کاربرد داده‌کاوی در پیش‌بینی مرگ بیماران سوختگی: مقایسه‌ی عملکرد چندین الگوریتم*

احسان نبوتی[1]، امیرعباس عزیزی[2]، ابراهیم عباسی[3]، حسن وکیلی ارکی[4]، جواد زارعی[5]، امیررضا رضوی[6]

مقاله پژوهشی

چکیده

مقدمه: در دهه‌ی اخیر الگوریتم‌های یادگیری ماشین به ابزار مفیدی جهت داده‌کاوی در داده‌های پزشکی، برای تولید مدل‌های پیش‌بینی تبدیل شده‌اند. سوختگی از جمله بیماری‌هایی است که پیش‌بینی پیامد آن اهمیت دارد. هدف این مطالعه تعیین عملکرد دو الگوریتم پراستفاده‌ی یادگیری ماشین یعنی شبکه‌ی عصبی مصنوعی و درخت تصمیم و مقایسه‌ی آن‌ها با روش آماری رگرسیون لجستیک در پیش‌بینی پیامد بیماران سوختگی بوده است.

روش بررسی: در این مطالعه مشاهده‌ای گذشته‌نگر پس از انجام پردازش اولیه‌ی داده‌ها و تعیین پیامد، دو الگوریتم یادگیری ماشین به همراه روش آماری رگرسیون لجستیک برای تولید مدل‌های پیش‌بینی روی داده‌های ۴۸۰۴ بیمار سوختگی مرکز آموزشی- درمانی آیت‌اله طالقانی شهر اهواز مربوط به سال‌های ۱۳۸۰ تا ۱۳۸۶ اعمال گردید. برای پردازش اولیه‌ی داده‌ها نرم‌افزار SPSS نسخه‌ی ۱۶ و در مرحله‌ی مدل‌سازی از Clementine 12.0 استفاده شد. همچنین با بکارگیری تکنیک 10-Fold Cross Validation، معیارهای ارزیابی کارایی برای داده‌های تست محاسبه و مقایسه شدند.

یافته‌ها: نتایج نشان داد الگوریتم شبکه‌ی عصبی مصنوعی با دقت ۹۷ درصد منجر به دقیق‌ترین مدل روی داده‌های مورد مطالعه می‌شود. مدل درخت تصمیم با دقت ۹۵ درصد در رده‌ی دوم و مدل رگرسیون لجستیک با دقت ۹۰ درصد کم‌ترین دقت را داشت. سایر معیارهای ارزیابی مانند حساسیت، ویژگی، PPV (Positive Predictive Value) و NPV (Negative Predictive Value) و AUC (Area Under the Curve) نیز کارایی مدل شبکه‌ی عصبی مصنوعی را بالاتر از دو مدل دیگر نشان دادند.

نتیجه‌گیری: تحلیل نتایج این مطالعه و مطالعات مشابه نشان می‌دهند که الگوریتم‌های یادگیری ماشین نسبت به روش‌های آماری منجر به تولید مدل‌های دقیق‌تری می‌شوند. بسته به ماهیت و میزان داده‌ها و همچنین جامعه‌ی پژوهش، الگوریتم‌های مختلف یادگیری ماشین، رفتارهای متفاوتی دارند که بنظر می‌رسد دقت مدل‌های شبکه‌ی عصبی مصنوعی از سایر مدل‌ها بیش‌تر می‌باشد.

واژه‌های کلیدی: داده‌کاوی؛ یادگیری ماشین؛ پیش‌بینی؛ درخت تصمیم؛ شبکه‌ی عصبی مصنوعی؛ سوختگی‌ها

دریافت مقاله: ۹۱/۹/۱۱ اصلاح نهایی: ۹۲/۳/۱۲

پذیرش مقاله: ۹۲/۵/۱۴

ارجاع: نبوتی احسان، عزیزی امیرعباس، عباسی ابراهیم، وکیلی ارکی حسن، زارعی جواد، رضوی امیررضا. **کاربرد داده‌کاوی در پیش‌بینی مرگ بیماران سوختگی: مقایسه عملکرد چندین الگوریتم.** مدیریت اطلاعات سلامت ۱۳۹۲؛ ۱۰(۶): ۷۹۹-۷۸۹.

مقدمه

وقوع حادثه‌ی سوختگی می‌تواند فاجعه‌ای غم‌انگیز برای فرد، خانواده و جامعه رقم بزند و سبب تحمیل صدمات جسمی، روحی- روانی، اقتصادی و اجتماعی بر آن‌ها گردد. سال‌هاست سوختگی از عوامل اصلی ناتوانی‌ها و مرگ و میر در جهان است،

* این مقاله حاصل تحقیق مستقل بدون حمایت مالی می‌باشد.

۱- دانشجوی دکتری تخصصی، انفورماتیک پزشکی، کمیته تحقیقات دانشجویی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی مشهد، مشهد و مربی، گروه مدیریت و فن‌آوری اطلاعات سلامت، دانشکده‌ی پیراپزشکی، دانشگاه علوم پزشکی کاشان، کاشان، ایران

۲- دانشجوی دکتری تخصصی، انفورماتیک پزشکی، کمیته تحقیقات دانشجویی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی مشهد، مشهد و مربی، گروه مدیریت و فن‌آوری اطلاعات سلامت، دانشکده‌ی پیراپزشکی، دانشگاه علوم پزشکی جندی شاپور اهواز، اهواز، ایران

۳- دانشجوی دکتری تخصصی، انفورماتیک پزشکی، کمیته تحقیقات دانشجویی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی مشهد، مشهد و پژوهشگر گروه پژوهشی انفورماتیک سرطان، مرکز تحقیقات سرطان پستان جهاد دانشگاهی، تهران، ایران

۴- دانشجوی دکتری تخصصی، انفورماتیک پزشکی، کمیته تحقیقات دانشجویی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران

۵- دانشجوی دکتری تخصصی، مدیریت اطلاعات سلامت، مرکز تحقیقات علوم مدیریت و اقتصاد سلامت، دانشکده‌ی مدیریت و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی ایران، تهران، ایران

۶- استادیار، گروه انفورماتیک پزشکی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران (نویسنده‌ی مسؤول) Email: razaviar@mums.ac.ir

به‌طوری‌که عامل ۵ تا ۱۲ درصد سوانح در جهان معرفی شده است (۱). سالیانه نزدیک به دو میلیون نفر در ایالات متحده دچار سوختگی می‌شوند (۲) و این حادثه پنجمین علت مرگ ناشی از حوادث در این کشور می‌باشد (۳). سوختگی همچنین درصد بالایی از حوادث را در ایران به خود اختصاص داده است (۴).

پیش‌بینی پیامد بیماری سوختگی براساس شرایط بیمار از اهمیت بالایی برخوردار است. از آنجایی‌که پیش‌بینی صحیح پیامد بیماری‌ها ابزار مناسبی برای متخصصین جهت تصمیم‌گیری بالینی آگاهانه به‌حساب می‌آید، پژوهشگران پزشکی همیشه به دنبال مدل‌هایی جهت پیش‌بینی پیامد بیماری‌ها و تعیین اهمیت فاکتورهای خطر بوده‌اند. پیش‌بینی صحیح علاوه بر اینکه نقش مؤثری در ادامه‌ی سیر درمان ایفا می‌کند، تأثیر قابل توجهی در تخصیص بهینه‌ی منابع بیمارستانی داشته و در نتیجه باعث ارتقای کیفیت ارایه‌ی خدمات مراقبت می‌گردد.

سال‌هاست پژوهشگران پزشکی از روش‌های متفاوت آماری برای مطالعه‌ی رابطه‌ی بین وضعیت بیماران و پیامد بیماری استفاده می‌کنند. از معروف‌ترین این مدل‌ها در زمینه‌ی پیش‌بینی پیامد بیماری سوختگی، مدل‌های Baux و ABSI می‌باشند (۵، ۶). تفاوت چنین مدل‌هایی در جمعیت مطالعه، متغیرهای پیش‌بینی، روش آماری و همچنین دقت پیش‌بینی می‌باشد. Baux از معروف‌ترین مدل‌های پیش‌بینی پیامد سوختگی است که از حاصل جمع سن با سطح سوختگی، احتمال فوت بیمار را مشخص می‌کند (۵). ABSI نیز یک سیستم امتیازدهی ساده و از لحاظ بالینی مفید می‌باشد که با استفاده از روش آماری رگرسیون لجستیک و متغیرهای گوناگون مانند سن، جنسیت، سطح سوختگی و درجه‌ی سوختگی، پیامد سوختگی را پیش‌بینی می‌کند (۶). اما مدل‌های مبتنی بر روش‌های آماری از دقت قابل قبولی برخوردار نبوده و با افزایش حجم داده‌ها، بهبود قابل توجهی در دقت آن‌ها مشاهده نمی‌شود.

در دهه‌ی اخیر با تسهیل جمع‌آوری داده‌های پزشکی و دسترس‌پذیری حجم زیاد داده‌های الکترونیکی، اکتشاف دانش توسعه یافته است. امروزه پژوهشگران از طریق فرایند اکتشاف دانش، اقدام به شناسایی روابط بین متغیرها، شناسایی الگوها و تولید مدل‌های پیش‌بینی پیامد بیماری‌ها می‌کنند. الگوریتم‌های یادگیری ماشین از ابزارهای مرسوم در فرایند اکتشاف دانش بوده که با استفاده از آن‌ها مدل‌های پیش‌بینی با دقت بالایی ایجاد می‌شوند. به استفاده از الگوریتم‌های یادگیری ماشین برای اکتشاف دانش در حجم عظیمی از داده‌ها، داده‌کاوی گفته می‌شود. مطالعات گوناگونی مزایای استفاده از داده‌کاوی در زمینه‌ی پزشکی را برشمرده‌اند (۱۰-۷). مطالعه‌ای که در سال ۲۰۱۰ منتشر شده به بررسی کاربرد داده‌کاوی در شناسایی فاکتورهای مهم در افتادن بیماران پرداخته است (۱۱). پژوهشگران برای ایجاد مدل‌های پیش‌بینی از الگوریتم شبکه‌ی عصبی مصنوعی ANN (Artificial Neural Network) و رگرسیون لجستیک و مجموعه داده‌ای مربوط به ۷۲۵ بیمار حاوی ۱۰ متغیر پیش‌بینی استفاده نمودند. یافته‌های پژوهشگران نشان داد که مدل حاصل از ANN با دقت ۷۷ درصد، نسبت به مدل رگرسیون لجستیک با دقت ۴۲ درصد، دقت بالاتری دارد. کریمی و همکاران در مطالعه‌ی خود به بررسی نقش ANN در پیش‌بینی بقای بیماران سوختگی پرداخته‌اند (۱۲). داده‌های این مطالعه مربوط به ۱۰۸۲ بیمار بود که به مرکز سوختگی بیمارستان شهید مطهری تهران در فاصله‌ی یک سال مراجعه کرده‌اند. در این مطالعه از الگوریتم ANN برای تولید مدل پیش‌بینی با به‌کارگیری ۱۵ متغیر، استفاده شده است. نویسندگان ۷۵ درصد از داده‌ها را برای آموزش مدل و ۲۵ درصد باقیمانده را برای تست آن استفاده نمودند. پس از ایجاد مدل، دقت بالای ۹۰ درصد برای این مدل پیشنهادی، گزارش شده است. به‌طور کلی مدل‌های تولید شده با تکنیک‌های داده‌کاوی از دقت بالایی برخوردار بوده و با افزایش حجم داده‌ها، پیش‌بینی‌های دقیق‌تری را برای نمونه‌های جدید انجام می‌دهند.

اما کدامیک از الگوریتم‌های داده‌کاوی دقت بالاتری در پیش‌بینی پیامد بیماری‌ها دارند؟ تاکنون مطالعات متعددی اقدام به مقایسه‌ی عملکرد انواع الگوریتم‌های یادگیری ماشین در پیش‌بینی پیامد بیماری‌ها کرده‌اند (۱۶-۱۳). مطالعه‌ای در سال ۲۰۰۴ میلادی عملکرد دو الگوریتم ANN و درخت تصمیم

به‌همراه روش رگرسیون لجستیک در پیش‌بینی بیماری سرطان پستان را با یکدیگر مقایسه نموده است (۱۳). از نقاط قوت آن مطالعه می‌توان به تعداد بالای رکوردها (بیش از ۲۰۰ هزار رکورد) برای آموزش مدل‌ها اشاره کرد. هر کدام از رکوردها حاوی ۱۶ متغیر پیش‌بینی کننده می‌باشد. یافته‌های پژوهشگران نشان داد که درخت تصمیم با دقت ۹۳/۶ درصد بالاترین دقت، ANN با ۹۱/۲ درصد در رده‌ی دوم و مدل رگرسیون لجستیک با ۸۹/۲ درصد پایین‌ترین دقت را داشته است. در مطالعه‌ای مشابه که در سال ۲۰۰۹ میلادی در کشور هندوستان انجام شده است، دقت چهار الگوریتم یادگیری ماشین (Navie Bayes، درخت تصمیم، SVM (Support Vector Machine) و ANN) در پیش‌بینی زنده ماندن بیماران سوختگی با هم مقایسه شده است (۱۴). در این پژوهش داده‌های ۱۸۰ بیمار سوختگی با فیلدهای داده‌ای سن، جنسیت و درصد سوختگی هشت بخش متفاوت بدن، برای ایجاد و تست مدل‌های پیش‌بینی استفاده شده است. براساس نتایج این مطالعه مدل Navie Bayes با دقت ۹۷/۷۸، بالاترین دقت، مدل‌های درخت تصمیم و SVM با دقت ۹۶/۱۲ در رده‌ی دوم و روش ANN با دقت ۹۵ درصد پایین‌ترین دقت را داشته است.

اما برای تولید مدل پیش‌بینی جدید براساس داده‌های بیماران سوختگی در کشور ایران استفاده از کدام الگوریتم توصیه می‌شود؟ آیا همیشه استفاده از یک الگوریتم منجر به دقیق‌ترین مدل‌ها می‌شود و یا اینکه با تغییر جامعه‌ی پژوهش، نوع و مقادیر متغیرهای پیش‌بینی کننده، دقت الگوریتم‌ها متفاوت می‌شود؟ براساس بررسی‌های انجام شده تاکنون در کشور ایران هیچ مطالعه‌ای عملکرد الگوریتم‌های داده‌کاوی در پیش‌بینی پیامد بیماری سوختگی را مقایسه نکرده است. در این مقاله هدف مقایسه‌ی الگوریتم‌های داده‌کاوی متفاوت جهت تولید مدل‌های پیش‌بینی مرگ بیماران سوختگی، در جامعه‌ی بیماران سوختگی در کشور ایران بوده است. برای این هدف از دو الگوریتم پرکاربرد یادگیری ماشین یعنی ANN و درخت تصمیم و همچنین روش آماری رگرسیون لجستیک برای تولید مدل‌های پیش‌بینی استفاده گردید. همچنین با استفاده از معیارهای ارزیابی شامل دقت، حساسیت (Sensitivity)، ویژگی (Specificity)، PPV، NPV و سطح زیر منحنی ROC (Receiver Operating Characteristic) کارایی مدل‌ها مورد ارزیابی قرار گرفت.

روش بررسی

منبع داده‌ها: مجموعه داده‌ی مورد استفاده در این پژوهش مربوط به بیماران سوختگی بستری شده‌ی مرکز آموزشی-درمانی آیت‌اله طالقانی شهر اهواز در فاصله‌ی زمانی سال‌های ۱۳۸۰ تا ۱۳۸۶ خورشیدی بوده است. این داده‌ها با استفاده از چک‌لیستی پژوهشگر ساخته که مورد تأیید متخصصین سوختگی بود، از پرونده‌ی ۶۰۲۴ پرونده‌ی بیمار سوختگی به‌صورت مقطعی و گذشته‌نگر استخراج شده است.

شناخت داده‌ها و پردازش اولیه (Data Understanding and Preprocessing): متغیرهای استخراج شده از پرونده‌ی بیماران شامل ۱۱ متغیر می‌باشد (جدول ۱ و ۲). تعداد رکوردهای استخراج شده از پرونده‌ها در مرحله‌ی اولیه ۶۰۲۴ بود. در مرحله‌ی پردازش اولیه، متغیر کد حذف و متغیر نتیجه‌ی درمان با دو مقدار زنده و فوت به‌عنوان پیامد در نظر گرفته شد. با حذف رکوردهایی که پیامد مشخصی نداشته و یا به میل شخصی مرخص شده بودند، تعداد رکوردها، جهت مدل‌سازی به ۴۸۰۴ رکورد کاهش یافت. جدول ۱ توزیع مربوط به متغیر وابسته‌ی پیامد را نشان می‌دهد. جدول ۲ نیز خلاصه‌ای از ۱۰ متغیر مستقل و پیش‌بینی کننده را نشان می‌دهد.

جدول ۱: توزیع متغیر وابسته‌ی پیامد

نتیجه‌ی درمان (پیامد)	تعداد	درصد
زنده (۱)	۳۴۱۶	۷۱/۱۱
فوت (۲)	۱۳۸۸	۲۸/۸۹

جدول ۲: جدول متغیرهای مستقل مدل‌های پیش‌بینی

نام متغیر گسسته	تعداد مقادیر منحصر به‌فرد	مقادیر
جنسیت	۲	مرد، زن
درجه‌ی سوختگی	۴	یک، دو، سه، هیچکدام
علت سوختگی	۶	مایعات و غذاهای داغ، سوختگی الکتریکی، آتش، قیر، مواد شیمیایی، فلزات داغ و مذاب
ماه پذیرش	۱۲	ماه‌های سال
کشت ادرار	۱۱	بدون کشت، کشت منفی، استافیلوک (Staphylococcus)، سودوموناس (Pseudomonas)، انتروباکتر (Enterobacter)، اشرشیاکولای (E-Coli)، کلبسیلا (Klebsialla)، باسیلوس پروتئوس (Bacillus Proteus)، استرپتوکوک (Streptococcal)، استینوباکتر (Actinobacter)، کشت مثبت غیر باکتریایی
کشت خون	۱۱	بدون کشت، کشت منفی، استافیلوک، سودوموناس، انتروباکتر، اشرشیاکولای، کلبسیلا، باسیلوس پروتئوس، استرپتوکوک، استینوباکتر، کشت مثبت غیر باکتریایی
کشت زخم	۱۱	بدون کشت، کشت منفی، استافیلوک، سودوموناس، انتروباکتر، اشرشیاکولای، کلبسیلا، باسیلوس پروتئوس، استرپتوکوک، استینوباکتر، کشت مثبت غیر باکتریایی

نام متغیر پیوسته	میانگین	S.D.	محدوده
سن (سال)	۲۳/۲۲	۱۶/۹۹	۱-۹۱
درصد سوختگی	۴۰/۷۹	۳۰/۸۵	۰/۵-۱۰۰
مدت بستری (روز)	۱۴/۶۹	۱۴/۴۷	۱-۱۰۰

ابتدا باید با استفاده از تکنیکی مجموعه‌ی داده‌ها را به زیرمجموعه‌هایی جداگانه برای ایجاد و تست مدل‌ها تفکیک کنیم. برای کاهش بایاس در مدل‌سازی استفاده از روش K-Fold Cross Validation برای تفکیک داده‌ها توصیه شده است (۱۹-۱۷). سه الگوریتم ANN، درخت تصمیم و رگرسیون لجستیک روی زیرمجموعه‌ی آموزش، اعمال و مدل‌های پیش‌بینی ایجاد می‌گردد. سپس مدل‌های ایجاد شده روی زیرمجموعه‌ی باقیمانده، تست و کارایی آن‌ها با استفاده از معیارهای حساسیت، ویژگی، دقت، PPV، NPV و AUC ارزیابی می‌شود. همچنین اهمیت هر کدام از متغیرها در پیش‌بینی پیامد مشخص می‌گردد.

الگوریتم‌های تولید مدل‌های پیش‌بینی: الگوریتم ANN به‌علت تولید مدل‌هایی با دقت بالا و الگوریتم‌های درخت تصمیم به‌علت سادگی و فهم آسان‌تر محبوبیت بیش‌تری نسبت به سایر الگوریتم‌های یادگیری ماشین دارند. به همین دلیل در مطالعات از این دو الگوریتم بیش‌تر از سایر الگوریتم‌ها استفاده می‌گردد. در این مطالعه این دو الگوریتم و روش رگرسیون لجستیک مقایسه می‌گردند.

ANN، ساختاری از نورون‌های مصنوعی است که به روشی نظام‌مند سازماندهی شده‌اند. نورون‌ها در چنین شبکه‌ای در قالب لایه‌های چندگانه (Multi-Layer Perception) سازماندهی می‌شوند. هر ANN حداقل از سه لایه شامل لایه‌های ورودی، میانی و خروجی تشکیل شده است. نورون‌های لایه‌ی میانی به‌صورت کامل به نورون‌های دو لایه‌ی دیگر متصل می‌شوند. هر کدام از اتصالات بین نورون‌ها دارای وزن می‌باشد که مقدار این وزن قدرت تأثیر نورون را روی نورون‌های بعدی نشان می‌دهد. مقادیر متغیرهای مستقل به نورون‌های لایه‌ی ورودی متصل شده و مقدار متغیر وابسته‌ی پیامد در لایه‌ی خروجی مشخص می‌شود. جریان اطلاعات از لایه‌ی ورودی به سمت لایه‌های میانی حرکت و در نهایت منجر به تعیین مقدار خروجی در لایه‌ی خروجی می‌شود و مدل را آموزش می‌دهد (۲۰).

الگوریتم‌های درخت تصمیم از الگوریتم‌های دسته‌بندی

قدرت‌مندی هستند که با پیشرفت دانش داده‌کاوی معروف‌تر نیز شده‌اند. در این تکنیک نمونه‌ها به‌صورت بازگشتی در شاخه‌های درخت قرار می‌گیرند تا مدل ساختار درختی، جهت پیش‌بینی با دقت بالا شکل بگیرد. از پراستفاده‌ترین الگوریتم‌های درخت تصمیم می‌توان به C4.5، C5، CHAID (Chi-squared Automatic Interaction Detector) و C&R Tree (Classification and Regression) اشاره کرد. در این پژوهش از الگوریتم CHAID برای درخت تصمیم استفاده شده است. این الگوریتم جهت ساخت درخت تصمیم (Decision Tree Induction) تمامی مقادیر متغیرهای پیش‌بینی کننده‌ی بالقوه را بررسی می‌کند. از خصوصیات مهم این الگوریتم این است که درخت تولید شده الزاماً دودویی نبوده، تمامی انواع متغیرها را ارزیابی می‌کند و همچنین رفتار مناسبی با مقادیر Missing دارد (۲۱).

رگرسیون لجستیک، یک روش آماری پراستفاده و نوعی آنالیز رگرسیون است که برای پیش‌بینی پیامد یک متغیر با تعداد محدودی از مقادیر ممکن، براساس متغیرهای پیش‌بینی کننده‌ی مستقل استفاده می‌شود.

نرم‌افزارهای مورد استفاده جهت پردازش اولیه و مدل‌سازی: در مرحله‌ی پردازش اولیه‌ی داده‌ها از نرم‌افزار SPSS نسخه‌ی ۱۶ و جهت ایجاد مدل‌های پیش‌بینی و تست آن‌ها، نرم‌افزار Clementine نسخه‌ی ۱۲ به‌کار گرفته شد. برای محاسبه‌ی مقادیر معیارهای ارزیابی کارایی مدل‌ها در ۱۰ مرحله‌ی آموزش و تست، میانگین آن‌ها و همچنین نمودارهای مربوطه، از نرم‌افزار Excel نسخه‌ی ۲۰۱۰ استفاده شد.

معیارهای ارزیابی کارایی مدل‌های پیش‌بینی: کارایی مدل‌ها در پیش‌بینی صحیح مقدار پیامد، معیار مهمی در انتخاب مدل پیش‌بینی می‌باشد. برای ارزیابی کارایی مدل‌ها در پیش‌بینی صحیح پیامد، معیارهای حساسیت، ویژگی، دقت، PPV، NPV و AUC مورد استفاده قرار می‌گیرد. برای محاسبه‌ی مقادیر این معیارها از پارامترهای جدول ماتریس Confusion استفاده می‌شود (جدول ۳). برای هر روش مدل‌سازی این ماتریس به‌صورت جداگانه ایجاد می‌گردد.

جدول ۳: ماتریس Confusion

		Predicted	
		Death	Life
Actual	Death	TP (True Positive)	FN (False Negative)
	Life	FP (False Positive)	TN (True Negative)

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$

$$PPV = \frac{TP}{TP+FP}$$

$$NPV = \frac{TN}{TN+FN}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

دقت: بیان‌گر دقت مدل در پیش‌بینی صحیح پیامد می‌باشد.

حساسیت: مدل با این احتمال برای بیماری که فوت خواهد کرد، پیش‌بینی مرگ می‌کند.

ویژگی: مدل با این احتمال برای بیماری که بهبود خواهد یافت، پیش‌بینی بهبودی می‌کند.

PPV: احتمال این‌که بیماری که مدل برای او پیش‌بینی مرگ کرده، فوت کند برابر با PPV می‌باشد.

NPV: احتمال این‌که بیماری که مدل برای او پیش‌بینی بهبودی کرده، بهبود یابد برابر با NPV می‌باشد.

ROC Curve: برای مقایسه‌ی کارایی مدل‌های پیش‌بینی استفاده می‌گردد. در این منحنی TP روی محور عمودی و FP روی محور افقی قرار می‌گیرد. هر چه مساحت زیر منحنی ROC بیش‌تر باشد، دقت مدل پیش‌بینی در تفکیک صحیح مقادیر پیامد بیش‌تر خواهد بود (۲۲).

تکنیک K-Fold Cross Validation: برای ارزیابی دقیق‌تر کارایی یک مدل بهتر است مجموعه داده به زیرمجموعه‌های آموزش و تست تقسیم گردد. سپس مدل پیش‌بینی با مجموعه داده‌ی آموزش ایجاد و دقت آن با مجموعه داده‌ی تست ارزیابی گردد. اما چگونه رکوردها در دو مجموعه قرار بگیرند و دو مجموعه چه اندازه‌ای داشته باشند تا میزان بایاس کاهش یابد؟

برای کاهش میزان بایاس، منابع متعدد استفاده از تکنیک K-Fold Cross Validation را توصیه می‌کنند (۱۹-۱۷). در این تکنیک مجموعه داده‌ی اصلی به‌صورت تصادفی در K زیرگروه داده‌ای جدا قرار می‌گیرند. مدل پیش‌بینی در هر بار با K-1 زیرگروه ایجاد و با یک زیرگروه باقیمانده تست می‌شود. این فرایند آموزش و تست، K بار تکرار می‌گردد. دقت کلی مدل پیش‌بینی در این تکنیک از معادله‌ی زیر محاسبه می‌شود.

$$CVA = \frac{1}{K}\sum_{i=1}^{K} Ai$$

در این معادله، CVA بیانگر میانگین دقت مدل‌های پیش‌بینی و A معیار ارزیابی حساسیت، ویژگی، دقت، PPV و NPV می‌باشد. در این مطالعه مانند بسیاری مطالعات دیگر K برابر ۱۰ در نظر گرفته شده است.

یافته‌ها

در این مطالعه بعد از مرحله‌ی پردازش اولیه، ارتباط متغیرهای پیش‌بینی کننده با پیامد از طریق قابلیت Feature Selection در نرم‌افزار Clementine مورد ارزیابی قرار گرفت. نتایج نشان داد که تمامی ۱۰ متغیر انتخابی در پیش‌بینی پیامد نقش مؤثری دارند. در ادامه، مدل‌های پیش‌بینی با استفاده از الگوریتم‌های ANN، درخت تصمیم (CHAID) و رگرسیون لجستیک ایجاد و معیارهای ارزیابی کارایی محاسبه شد. جدول ۴ مقادیر معیارهای ارزیابی کارایی مدل‌های پیش‌بینی را در مرحله تست برای ۱۰ مرحله و همچنین میانگین مقادیر آن‌ها را نشان می‌دهد.

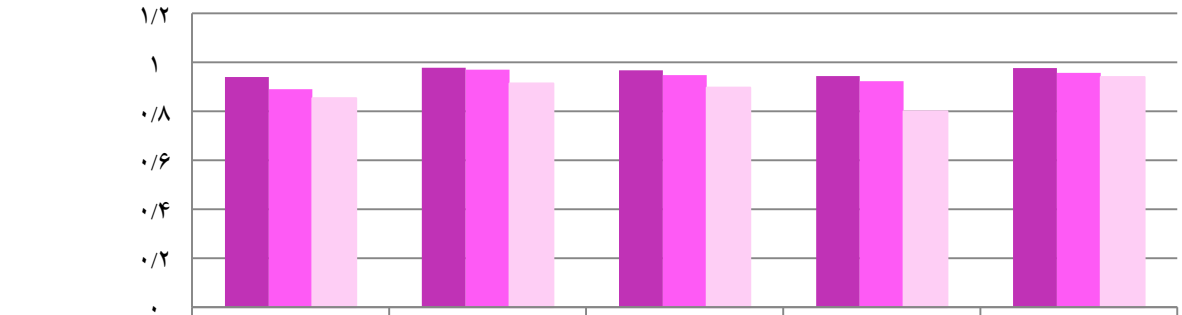
جدول ۴: مقادیر معیارهای ارزیابی کارایی سه مدل پیش‌بینی در مرحله‌ی تست

	رگرسیون لجستیک					درخت تصمیم (CHAID)					ANN				
	دقت	حساسیت	ویژگی	PPV	NPV	دقت	حساسیت	ویژگی	PPV	NPV	دقت	حساسیت	ویژگی	PPV	NPV
گروه داده‌ای ۱	۰/۸۹۹۳	۰/۸۵۵۱	۰/۹۰۹۱	۰/۷۸۶۷	۰/۹۴۱۲	۰/۹۲۵۹	۰/۹۷۱۶	۰/۹۵۳۱	۰/۹۶۳۴	۰/۹۷۱۶	۰/۹۷۹۶	۰/۹۷۱۰	۰/۹۸۳۰	۰/۹۵۷۱	۰/۹۸۸۶
گروه داده‌ای ۲	۰/۸۶۹۸	۰/۸۴۲۱	۰/۸۸۰۳	۰/۷۲۷۳	۰/۹۳۶۴	۰/۹۵۰۴	۰/۸۷۲۲	۰/۹۸۰۱	۰/۹۴۳۱	۰/۹۵۲۹	۰/۹۶۴۹	۰/۹۳۲۳	۰/۹۷۷۲	۰/۹۳۹۴	۰/۹۷۴۴
گروه داده‌ای ۳	۰/۸۸۹۸	۰/۸۵۰۷	۰/۹۰۵۳	۰/۷۸۰۸	۰/۹۳۸۷	۰/۹۴۷۰	۰/۸۹۵۵	۰/۹۶۷۵	۰/۹۱۶۰	۰/۹۵۸۹	۰/۹۶۴۰	۰/۹۴۷۸	۰/۹۷۰۴	۰/۹۲۷۰	۰/۹۷۹۱
گروه داده‌ای ۴	۰/۹۱۸۳	۰/۹۰۸۵	۰/۹۲۲۶	۰/۸۳۷۷	۰/۹۵۸۲	۰/۹۴۴۱	۰/۸۸۷۳	۰/۹۶۹۰	۰/۹۲۶۵	۰/۹۵۱۴	۰/۹۷۶۳	۰/۹۴۳۷	۰/۹۹۰۷	۰/۹۷۸۱	۰/۹۷۵۶
گروه داده‌ای ۵	۰/۸۹۵۳	۰/۸۲۸۶	۰/۹۲۰۸	۰/۸۰۰۰	۰/۹۳۳۵	۰/۹۴۸۶	۰/۸۷۸۶	۰/۹۷۵۴	۰/۹۳۱۸	۰/۹۵۴۵	۰/۹۸۰۲	۰/۹۵۰۰	۰/۹۹۱۸	۰/۹۷۷۹	۰/۹۸۱۱
گروه داده‌ای ۶	۰/۸۹۱۰	۰/۸۲۳۵	۰/۹۱۴۰	۰/۷۶۵۶	۰/۹۳۸۲	۰/۹۲۹۵	۰/۸۳۱۹	۰/۹۶۲۸	۰/۸۸۳۹	۰/۹۴۳۸	۰/۹۴۸۷	۰/۸۸۲۴	۰/۹۷۱۳	۰/۹۱۳۰	۰/۹۶۰۳
گروه داده‌ای ۷	۰/۹۱۰۲	۰/۸۶۲۷	۰/۹۳۲۵	۰/۸۵۷۱	۰/۹۳۵۴	۰/۹۵۶۲	۰/۹۰۲۰	۰/۹۸۱۶	۰/۹۵۸۳	۰/۹۵۵۲	۰/۹۶۸۷	۰/۹۳۴۶	۰/۹۸۴۷	۰/۹۶۶۲	۰/۹۶۹۸
گروه داده‌ای ۸	۰/۸۹۱۳	۰/۸۲۴۸	۰/۹۱۹۵	۰/۸۱۲۹	۰/۹۲۵۲	۰/۹۴۳۵	۰/۹۰۵۱	۰/۹۵۹۸	۰/۹۰۵۱	۰/۹۵۹۸	۰/۹۵۸۷	۰/۹۳۴۳	۰/۹۶۹۰	۰/۹۲۷۵	۰/۹۷۲۰
گروه داده‌ای ۹	۰/۹۰۲۰	۰/۸۶۳۰	۰/۹۱۸۶	۰/۸۱۸۲	۰/۹۴۰۵	۰/۹۵۱۰	۰/۹۱۷۸	۰/۹۶۵۱	۰/۹۱۷۸	۰/۹۶۵۱	۰/۹۵۵۱	۰/۹۵۸۹	۰/۹۵۳۵	۰/۸۹۷۴	۰/۹۸۲۰
گروه داده‌ای ۱۰	۰/۹۰۲۰	۰/۸۷۶۷	۰/۹۱۲۸	۰/۸۱۰۱	۰/۹۴۵۸	۰/۹۴۴۹	۰/۸۹۷۳	۰/۹۶۵۱	۰/۹۱۶۱	۰/۹۵۶۸	۰/۹۶۷۳	۰/۹۳۸۴	۰/۹۷۹۷	۰/۹۵۱۴	۰/۹۷۴۰
میانگین	۰/۸۹۶۸	۰/۸۵۳۶	۰/۹۱۳۶	۰/۷۹۹۶	۰/۹۳۹۳	۰/۹۴۶۸	۰/۸۸۹۳	۰/۹۶۹۸	۰/۹۲۲۵	۰/۹۵۶۲	۰/۹۶۶۴	۰/۹۳۹۳	۰/۹۷۷۱	۰/۹۴۳۵	۰/۹۷۵۷

برای مقایسه کارایی سه مدل در پیش‌بینی صحیح پیامد باید میانگین معیارهای ارزیابی کارایی آن‌ها در مرحله‌ی تست، حاصل از ۱۰ مرحله را با هم مقایسه کرد. نمودار ۱ میانگین مقادیر معیارها در مرحله‌ی تست را نشان می‌دهد.

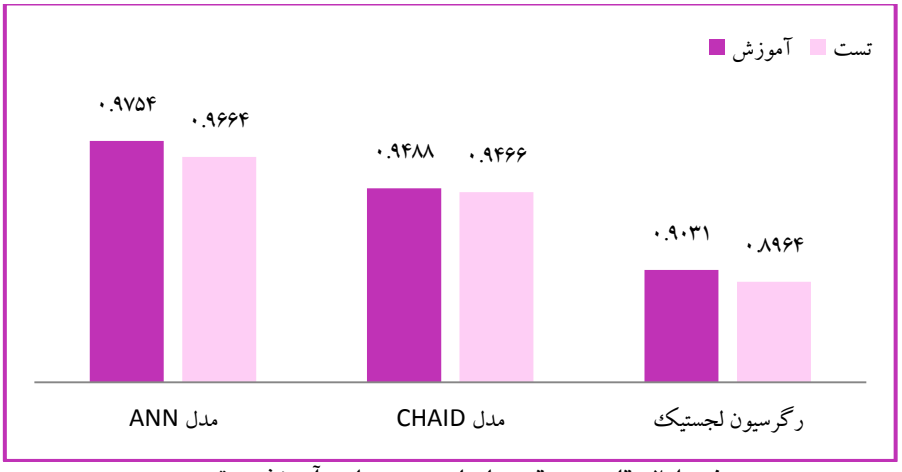
دقت مدل‌ها در دو مرحله‌ی آموزش و تست: نمودار ۲، دقت مدل‌ها در دو مرحله‌ی آموزش و تست را با یکدیگر مقایسه می‌کند.

اهمیت متغیرها در پیش‌بینی پیامد: از آنجایی‌که از ۱۰ متغیر متفاوت برای ایجاد مدل‌های پیش‌بینی استفاده شد، هرکدام از آن‌ها نقش متفاوتی در پیش‌بینی پیامد دارند. با انجام آنالیز حساسیت (Sensitivity Analysis)، مهمترین متغیر در پیش‌بینی پیامد در هر دو مدل ANN و درخت تصمیم درصد سوختگی بوده است و پس از آن متغیرهای درجه‌ی سوختگی، مدت زمان بستری، جنسیت و سن به‌ترتیب از اهمیت پایین‌تری برخوردار بوده‌اند. البته لازم به ذکر است که مدل ANN از تمامی ۱۰ متغیر پیش‌بینی، ولی مدل درخت تصمیم تنها از ۶ متغیر مهم‌تر استفاده کرده است.



	حساسیت	ویژگی	دقت	PPV	NPV
مدلANN	۰/۹۳۹۳	۰/۹۷۷۱	۰/۹۶۶۴	۰/۹۴۳۵	۰/۹۷۵۷
مدلCHAID	۰/۸۸۹۳	۰/۹۶۹۸	۰/۹۴۶۸	۰/۹۲۲۵	۰/۹۵۶۲
رگرسیون لجستیک	۰/۸۵۳۶	۰/۹۱۳۶	۰/۸۹۶۴	۰/۷۹۹۶	۰/۹۳۹۳

نمودار ۱: میانگین مقادیر معیارهای ارزیابی کارایی سه مدل پیش‌بینی



نمودار۲: مقایسه‌ی دقت مدل‌ها در دو مرحله‌ی آموزش و تست

بحث

با به‌کارگیری دو الگوریتم پرکاربرد یادگیری ماشین یعنی ANN و درخت تصمیم و همچنین روش آماری رگرسیون لجستیک روی داده‌های بیماران سوختگی مرکز آموزشی- درمانی آیت‌اله طالقانی اهواز، سه مدل پیش‌بینی پیامد بیماری ایجاد شد که مدل ANN بیش‌ترین دقت و مدل رگرسیون لجستیک کم‌ترین دقت را داشته است. همچنین مدل درخت تصمیم با به‌کارگیری شش متغیر پیش‌بینی از ۱۰ متغیر موجود، مدل ساده‌تری را با دقت قابل قبولی ایجاد کرد.

میزان داده‌ها: در منابع توصیه شده برای استفاده از الگوریتم‌های یادگیری ماشین جهت داده‌کاوی، میزان داده‌ها تا حد امکان زیاد باشد که با افزایش داده‌ها، دقت مدل‌ها افزایش می‌یابد (۱۹-۱۷). نتایج این مطالعه و مطالعات مشابه نیز این موضوع را تأیید می‌کند. در دو مطالعه‌ی مشابه با مطالعه‌ی حاضر، میزان داده‌ها به‌ترتیب ۱۰۸۲ و ۱۸۰ رکورد بود، در حالی‌که میزان داده‌ها در مطالعه‌ی حاضر ۴۸۰۴ می‌باشد (۱۴، ۱۲). بنابراین از نقاط قوت این مطالعه نسبت به مطالعات مشابه می‌توان به میزان زیاد داده‌های آن اشاره کرد که در نتیجه منجر به دقت بالاتری

در مدل ANN (۹۷ درصد) نسبت به دو مطالعه‌ی دیگر (۹۵ درصد و ۹۰ درصد) شده است.

مقایسه‌ی دقت مدل‌ها: همانگونه که در نمودار ۱ نشان داده شده است، کلیه‌ی معیارهای ارزیابی مدل ANN نسبت به مدل‌های دیگر بالاتر می‌باشد. مطالعه‌ی Patil و همکاران (۱۴) دقت مدل ANN را ۹۵ درصد و دقت مدل درخت تصمیم را ۹۶ درصد گزارش کرده که کم بودن دقت مدل ANN نسبت به درخت تصمیم احتمالاً ناشی از کم بودن میزان داده‌ها (۱۸۰ رکورد) و یا تفاوت ماهیت متغیرها در این مطالعه نسبت به مطالعه‌ی حاضر می‌باشد.

در مقایسه‌ی دقت الگوریتم‌های یادگیری ماشین با روش‌های آماری مانند رگرسیون لجستیک، مطالعات متعددی دقت الگوریتم‌های یادگیری ماشین را بالاتر گزارش کرده‌اند. در مطالعه‌ی Delen و همکاران (۱۳) که به پیش‌بینی پیامد بیماری سرطان پستان پرداخته، دقت مدل ANN و درخت تصمیم، بالاتر از مدل رگرسیون لجستیک اعلام شده است. همچنین در مطالعه‌ی Lee و همکاران (۱۱) که به مقایسه‌ی مدل‌های ANN و رگرسیون لجستیک در مورد پیش‌بینی آسیب بیماران بستری بعد از افتادن پرداخته، دقت مدل ANN بیش‌تر از مدل دیگر است. در مطالعه‌ی حاضر نیز دقت مدل ANN و درخت تصمیم (۹۷ درصد و ۹۵ درصد) از دقت رگرسیون لجستیک (۹۰ درصد) بیش‌تر است.

مقایسه‌ی سایر معیارهای ارزیابی مدل‌ها: نمودار ۱ نشان می‌دهد که علاوه بر دقت، سایر معیارهای ارزیابی مدل ANN نیز نسبت به دو مدل دیگر بیش‌تر می‌باشند. PPV برای مدل ANN برابر با ۰/۹۴۳۵ است، یعنی چنانچه این مدل برای بیماری پیش‌بینی مرگ کند، احتمال اینکه بیمار بمیرد ۹۴ درصد می‌باشد و PPV برابر با ۰/۷۹۹۶ برای مدل رگرسیون لجستیک نشان می‌دهد که چنانچه این مدل برای بیماری پیش‌بینی مرگ کند احتمال مرگ وی حدود ۸۰ درصد می‌باشد. بنابراین برای پیش‌بینی مرگ، مدل‌های ANN و درخت تصمیم نسبت به مدل رگرسیون لجستیک دقت بالاتری دارند. نتایج همچنین نشان می‌دهد که برای پیش‌بینی زنده ماندن، هر سه مدل با NPV بالای ۹۰ درصد مناسب هستند و از بین آن‌ها ANN بهترین است. همان‌گونه که قبلاً بیان شد، هر چه مساحت سطح زیر منحنی ROC بیش‌تر باشد، دقت مدل در تفکیک صحیح مقادیر پیامد بیش‌تر خواهد بود. مدل ANN با AUC برابر با ۰/۹۹۱۸ بالاترین قدرت تفکیک‌پذیری را بین مقادیر پیامد دارد. مدل‌های درخت تصمیم و رگرسیون لجستیک نیز با AUC بالای ۹۸ درصد قدرت تفکیک‌پذیری خوبی داشته‌اند.

تفکیک داده‌ها برای دو مرحله‌ی آموزش و تست: مسأله‌ی مهم در ارزیابی دقت مدل‌های پیش‌بینی، نحوه‌ی تفکیک داده‌ها به دو بخش جهت آموزش و تست مدل‌ها می‌باشد. مطالعه‌ی Estahbanati و Bouduhi (۱۲)، ۷۵ درصد از داده‌ها را برای آموزش مدل و ۲۵ درصد باقیمانده را برای تست آن به‌کار گرفته که از نقاط ضعف آن محسوب می‌گردد. در منابع تذکر داده شده است که با این روش تفکیک، احتمال وقوع بایاس در مدل‌سازی وجود دارد و برای تولید مدل‌های قابل اعتمادتر، استفاده از روش K-Fold Cross Validation با مقدار ۱۰ برای K توصیه شده است (۱۹-۱۷). در این مطالعه نیز مانند دو مطالعه‌ی Delen و همکاران و Patil و همکاران (۱۴، ۱۳) این تکنیک به کار گرفته شد تا بایاس در ایجاد مدل‌ها از بین برود و این از نقاط قوت مطالعه‌ی حاضر است.

مقایسه‌ی دقت مدل‌ها در دو مرحله‌ی آموزش و تست: همانند بسیاری از مدل‌های دیگر، دقت مدل‌های این مطالعه در مرحله‌ی تست کم‌تر از مرحله‌ی آموزش می‌باشد (نمودار ۲). دلیل آن این است که وقتی مدل با داده‌های آموزش تست شود، دقت بالاتری نسبت به وقتی که همان مدل با داده‌ای دیگر مورد ارزیابی قرار گیرد، دارد. البته نکته‌ی قابل توجه این است که تفاوت دقت سه مدل برای دو مرحله‌ی آموزش و تست مختصر است، که این به دلیل همگن بودن داده‌ها و کیفیت بالای آن‌ها می‌باشد. لازم به ذکر است که ناهمگنی داده‌ها، عملکرد الگوریتم‌های یادگیری ماشین را مختل و استفاده از آن‌ها را با مشکل مواجه می‌سازد.

بررسی اهمیت متغیرها در مدل‌های پیش‌بینی: همان‌گونه که قبلاً بیان شد، آنالیز حساسیت روشی برای استخراج رابطه‌ی بین متغیرهای مستقل و مقدار پیامد در مدل پیش‌بینی می‌باشد. نتایج آنالیز حساسیت نشان می‌دهند که مدل ANN از تمامی

۱۰ متغیر موجود برای پیش‌بینی پیامد استفاده کرده، در حالی‌که مدل CHAID از شش متغیر استفاده می‌کند. در مطالعه‌ی مشابه از دو متغیر سن و جنسیت به‌علاوه‌ی هشت متغیر مربوط به درصد سوختگی نواحی مختلف بدن برای پیش‌بینی پیامد استفاده شده است (۱۴). در مطالعه‌ی مذکور نیز الگوریتم درخت تصمیم تنها از شش متغیر استفاده کرده در حالی‌که الگوریتم ANN تمامی ۱۰ متغیر موجود را به‌کار گرفته است. از آنجایی‌که در تحلیل اولیه نشان داده شد که تمامی ۱۰ متغیر انتخابی در پیش‌بینی پیامد نقش دارند، بنابراین مدل ANN که از تمامی ۱۰ متغیر استفاده می‌کند مدل دقیق‌تری می‌باشد. از طرفی چون درخت تصمیم از متغیرهای کم‌تری استفاده می‌کند، بنابراین مدل ساده‌تری تولید خواهد کرد. مهم‌ترین متغیرها در پیش‌بینی پیامد در مطالعه‌ی حاضر برای هر دو مدل ANN و CHAID به ترتیب اهمیت، درصد سوختگی، درجه‌ی سوختگی، مدت زمان بستری، جنسیت و سن می‌باشند. در مطالعه‌ی Patil و همکاران (۱۴) نیز درصد سوختگی و سن بیمار مهمترین متغیرها در پیش‌بینی پیامد بوده‌اند. بنابراین به نظر می‌رسد که بهتر است متخصصین بالینی توجه بیش‌تری به مقادیر این پنج متغیر داشته باشند.

محدودیت‌ها: در حالیکه داده‌کاوی از طریق تولید مدل‌های پیش‌بینی دقیق توانایی فراهم‌سازی اطلاعات مفیدی را برای تصمیم‌گیران حوزه‌های بالینی دارد، محدودیت‌هایی نیز دارد. یکی از مشکلاتی که معمولاً باعث کاهش کیفیت داده‌ها می‌شود، وجود مقادیر Missing می‌باشد. در مطالعه‌ی حاضر تعداد رکوردهای اولیه ۶۰۲۴ مورد بود که به علت وجود مقادیر Missing در پیامد، این تعداد به ۴۸۰۴ مورد کاهش یافته است. واضح است که هر چه تعداد رکوردها کاهش می‌یابد، دقت مدل پیش‌بینی دستخوش تغییرات می‌شود. موضوع دیگر، پیچیدگی استفاده از مدل‌های داده‌کاوی برای پزشکان می‌باشد. گاهی اوقات مدل‌های داده‌کاوی به قدری پیچیده می‌شود (به‌خصوص ANN) که عملکرد آن‌ها برای پزشکان قابل فهم نمی‌باشد و در نتیجه ایشان در مورد استفاده از آن‌ها از خود مقاومت نشان می‌دهند. بنابراین یکی از پارامترهای پذیرش مدل‌ها توسط پزشکان، منطقی و قابل فهم بودن آن‌ها می‌باشد (مانند درخت تصمیم). نکته‌ی آخر اینکه استفاده از مدل‌های پیش‌بینی، هیچگاه جایگزینی برای تصمیم‌گیری‌های پزشکان نمی‌باشد، بلکه هدف، فراهم‌سازی اطلاعات تکمیلی و پشتیبان برای ایشان در اتخاذ تصمیمات آگاهانه‌تر است.

کاربرد مدل‌های پیشنهادی برای متخصصین بالینی: نتایج این مطالعه می‌تواند کاربرد زیادی برای متخصصین بالینی درگیر در درمان بیماری سوختگی داشته باشد. مدل‌های پیشنهادی می‌توانند با دقت بالای ۹۰ درصد پیامد بیمار سوختگی را پیش‌بینی کنند. پیش‌بینی صحیح مطمئناً منجر به استفاده‌ی بهینه از منابع موجود در درمان بیماران خواهد شد. همچنین آنالیز حساسیت، اهمیت متغیرهای پیش‌بینی در پیامد بیماری‌ها را نشان داد که این ترتیب اهمیت، پیامی به متخصصین بالینی در شناسایی متغیرهای مهم برای بیماران سوختگی می‌باشد.

نتیجه‌گیری

در این مطالعه با استفاده از سه الگوریتم ANN، درخت تصمیم و رگرسیون لجستیک اقدام به ایجاد سه مدل پیش‌بینی پیامد بیماران سوختگی شد و با تست مدل‌ها، معیارهای ارزیابی کارایی محاسبه و سه مدل با یکدیگر مقایسه گردید. نتایج نشان می‌دهد که الگوریتم‌های یادگیری ماشین نسبت به روش‌های آماری منجر به مدل‌های دقیقتری می‌شوند که در این میان دقت ANN از سایر روش‌ها بیش‌تر می‌باشد، در حالی‌که درخت تصمیم با استفاده از متغیرهای کم‌تر، مدل‌های ساده‌تر و قابل فهمی را ایجاد می‌کند. در مقایسه با مطالعات مشابه نتیجه‌گیری شد که براساس نوع متغیرها، ماهیت و میزان داده‌ها و همچنین جامعه‌ی پژوهش، الگوریتم‌های مختلف رفتارهای متفاوتی نشان می‌دهند. بنابراین پیشنهاد می‌گردد که برای هر محیطی متناسب با شرایط موجود، اقدام به بررسی کارایی الگوریتم‌های متفاوت شود تا بدین ترتیب مناسب‌ترین الگوریتم برای تولید مدل پیش‌بینی انتخاب گردد. شاید هیچ‌گاه نتوان مدلی تولید کرد که به‌صورت قطعی قادر به پیش‌بینی پیامد تمامی بیماران سوختگی باشد، ولی با به‌کارگیری الگوریتم‌های متعدد داده‌کاوی، جمع‌آوری هدفمند داده‌ها، پردازش اولیه‌ی مناسب و ارتقای کیفیت داده‌ها می‌توان مدل‌های پیش‌بینی با دقت قابل قبولی را ارایه کرد.

پیشنهادها

این مطالعه تنها براساس داده‌های یک مرکز سوختگی انجام شده است. پیشنهاد می‌شود همین مطالعه براساس داده‌های سایر مراکز سوختگی نیز تکرار گردد و با مقایسه‌ی نتایج آن با نتایج مطالعه‌ی حاضر، در مورد دقت مدل‌های پیش‌بینی برای بیماران سوختگی کشور ایران با اطمینان بیش‌تری بحث کرد. مدل‌های پیشنهادی در این مطالعه تنها براساس داده‌های بیماران و بدون استفاده از دانش متخصصین سوختگی ایجاد شده‌اند؛ پیشنهاد می‌شود در مطالعات آینده علاوه بر داده‌های بیماران، از دانش و تجربه‌ی متخصصین سوختگی نیز از طریق روش‌هایی مانند منطق فازی استفاده گردد.

تشکر و قدردانی

بدینوسیله نویسندگان مقاله، مراتب سپاس و قدردانی خود را از پرسنل بخش مدارک پزشکی مرکز آموزشی- درمانی آیت‌اله طالقانی اهواز، به خاطر مشارکت در گردآوری داده‌ها، اعلام می‌دارند.

References

1. Olaitan PB, Olaitan JO. Burns and scalds--epidemiology and prevention in a developing country. Niger J Med 2005; 14(1):9-16.
2. Herndon DN, Spies M. Modern burn care. Semin Pediatr Surg 2001; 10(1): 28-31.
3. Badger JM. Burns: the psychological aspects. Am J Nurs 2001; 101(11): 38-44.
4. Afrasiabifar A, Karimi Z. Causes and materials of burning among the patients hospitalized in Yasuj Shahid Beheshti Hospital 2001. Armaghane Danesh 2002; 7(27): 39-46. [In Persian]
5. Osler T, Glance LG, Hosmer DW. Simplified estimates of the probability of death after burn injuries: extending and updating the baux score. J Trauma 2010; 68(3): 690-7.
6. Tobiasen J, Hiebert JM, Edlich RF. The abbreviated burn severity index. Ann Emerg Med 1982; 11(5): 260-2.
7. Hobbs GR. Data mining and healthcare informatics. Am J Health Behav 2001; 25(3): 285-9.
8. Yoo I, Alafaireet P, Marinov M, Pena-Hernandez K, Gopidi R, Chang JF, et al. Data mining in healthcare and biomedicine: a survey of the literature. J Med Syst 2012; 36(4): 2431-48.
9. Obenshain MK. Application of data mining techniques to healthcare data. Infect Control Hosp Epidemiol 2004; 25(8): 690-5.
10. Koh HC, Tan G. Data mining applications in healthcare. J Healthc Inf Manag 2005; 19(2): 64-72.
11. Lee TT, Liu CY, Kuo YH, Mills ME, Fong JG, Hung C. Application of data mining to the identification of critical factors in patient falls using a web-based reporting system. Int J Med Inform 2011; 80(2): 141-50.
12. Estahbanati HK ,Bouduhi N. Role of artificial neural networks in prediction of survival of burn patients-a new approach. Burns 2002; 28(6): 579-86.
13. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. Artif Intell Med 2005; 34(2):113-27.
14. Patil BM, Joshi RC, Toshniwal D, Biradar S. A new approach: role of data mining in prediction of survival of burn patients. J Med Syst 2011; 35(6): 1531-42.
15. Oztekin A, Delen D, Kong ZJ. Predicting the graft survival for heart-lung transplantation patients: an integrated data mining methodology. Int J Med Inform 2009; 78(12): e84-96.
16. Chu CM, Chang YT, Wetter T. Artificial neural network may perform good to predict the survivability of cervical cancer. AMIA Annu Symp Proc. United States 2006; 2006: 889.
17. Cios KJ, Pedrycz W, Swiniarski RW, Kurgan LA. Data Mining: A Knowledge Discovery Approach. USA: Springer; 2007: 473-4.
18. Olson DL, Delen D. Advanced Data Mining Techniques. USA: Springer; 2008: 141:3.
19. Witten IH, Frank E, Hall MA. Data Mining: Practical Machine Learning Tools and Techniques. USA: Esevier; 2005: 149-51.
20. Larose DT .Discovering Knowledge in Data: An Introduction to Data Mining. Canada: John Wiley and Sons; 2005: 128:45.
21. Clementine 12.0 Algorithms Guide. USA: SPSS Inc; 2007: 43-55.
22. Shortliffe EH, Cimino JJ. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. USA: Springer; 2006: 91-4.

Using Data Mining to Predict Outcome in Burn Patients: A Comparison between Several Algorithms*

Ehsan Nabovati[1]; Amir Abbas Azizi[2]; Ebrahim Abbasi[3]; Hassan Vakili-Arki[4]; Javad Zarei[5]; Amir Reza Razavi[6]

Original Article

Abstract

Introduction: In the past decades, machine learning algorithms have become a useful tool for data mining within huge amounts of health data to create prediction models. Burn is one of the diseases that predicting of its outcome has high importance. The aim of this study was to survey two widely used machine learning algorithms; Artificial Neural Network (ANN) and decision tree, and compare them with logistic regression method to predict the outcome of burn patients.

Methods: In this retrospective observational study, following preprocessing of the data and determining the outcome of patient (Life or death), two well-known machine learning algorithms (ANN and decision tree) and logistic regression method were used to create prediction models using data from 4804 burn patients hospitalized in Taleghani Burn Center in Ahvaz during the years 2001-2007. The preprocessing of the data was performed using SPSS (Version 16.0), and in the modeling phase, Clementine (Version 12.0) software was used. Moreover, 10-fold cross validation technique was used to validate the model and criteria for evaluating the performance of models were measured and compared.

Results: The results showed that the ANN algorithm with accuracy of 97% resulted the most accurate model on the studied data. The decision tree model with 95% accuracy was in the second place and the logistic regression model with an accuracy of 90% was the least accurate. Moreover other evaluating criteria such as sensitivity, specificity, PPV, NPV and AUC showed that performance of the ANN model was better than the others.

Conclusion: The current study shows that machine learning algorithms compared with statistical methods create more accurate models. In analyzing the current data, the model created by ANN is more accurate than the other machine learning algorithm, decision tree.

Keywords: Data Mining; Machine Learning; Forecasting; Decision Tree; Artificial Neural Network; Burns

Received: 1 Dec, 2012 Accepted: 5 Aug, 2013

Citation: Nabovati E, Azizi AA, Abbasi E, Vakili-Arki H, Zarei J, Razavi AR. **Using Data Mining to Predict Outcome in Burn Patients: A Comparison between Several Algorithms.** Health Inf Manage 2014; 10(6):799.

* This article was an independent research with no financial aid.

1- PhD Candidate, Medical Informatics, Student Research Committee, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad and Lecturer, Management and Health Information Technology Department, Faculty of Paramedicine, Kashan University of Medical Sciences, Kashan, Iran

2- PhD Candidate, Medical Informatics, Student Research Committee, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad and Lecturer, Management and Health Information Technology Department, Faculty of Paramedicine, Ahwaz Jundishapur University of Medical Sciences, Ahwaz, Iran

3- PhD Candidate, Medical Informatics, Student Research Committee, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad and Researcher of Cancer Informatics Research Group, BCRC(Breast Cancer Research Center), ACECR (Academic Center for Education, Culture and Research), Tehran, Iran

4- PhD Candidate, Medical Informatics, Student Research Committee, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

5- PhD Candidate, Health Information Management, Health Management and Economics Research Center, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran

6- Assistant Professor, Department of Medical Informatics, Mashhad University of Medical Sciences, Mashhad, Iran (Corresponding Author) Email: razaviar@mums.ac.ir